Generic and configurable diagnosis function based on production data stored in Manufacturing Execution System

N-H Tran, Henry Sébastien, Eric Zamaï

Generic and configurable diagnosis function based on production data stored in Manufacturing Execution System

N-H.Tran[1], S.Henry[2], and E.Zamaï[3]

[1,3] *G-SCOP Laboratory, CNRS UMR 5272, Grenoble-INP, France*
Ngoc-Hoang.Tran@grenoble-inp.fr
Eric.zamai@grenoble-inp.fr

[2] *DISP Laboratory, University Lyon 1, France*
Sebastien.Henry@univ-lyon1.fr

ABSTRACT

The paper proposes a diagnosis approach corresponding to the specific MES level to provide information on the origins of a performance indicator degradation. Our key distribution is the proposal of a set of potential causes that may impact the successful completion of production operations, such as the operator stress, quality of material, equipment or recipe change and their characteristic parameters by exploiting MES historical database. We use Bayesian Network model to diagnose the potential failure causes and support effective human decisions on corrective actions (maintenance, human resource planning, recipe re-qualification, etc) by computing conditional probabilities for each suspected proposed causes.

1. INTRODUCTION

Nowadays in a highly competitive and complex production environment with wide range products, manufacturers must be equipped with precise knowledge of the production systems via data analysis tools to support their process control. The information system like Manufacturing Execution System (MES) is designed for this requirement. MES is as a bridge between the Enterprise Resource Planning (ERP) and the local control, integrating several functions to control a production system based today on a standard data base (IEC, 2003). The MES solutions editor offers generic functions such as recipe management, execution production, traceability or performance analysis. The genericity is based on parametered functions that guarantee a cheaper and faster deployment. MES solutions collect and record a growing number of production data especially with the development of unitary traceability (as Fig. 1).

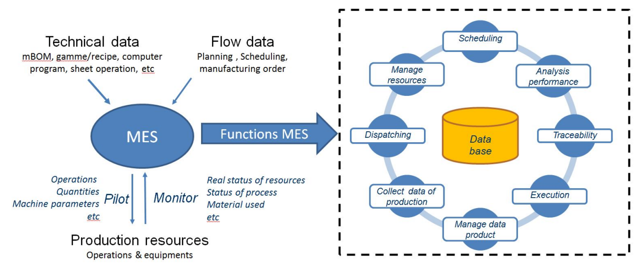


Figure 1. Manufacturing Execution System.

However, the exploitation of MES historical data is often limited to the calculations of Key Performance Indicators (KPI), such as Overall Equipment Effectiveness (OEE). Today, the evolution analysis of these indicators is performed by the users (operator, team leader, manager and directors) based on their knowledge on the production system and especially their expertise. In the complex and high variability context, the required time for these manual analysis becomes incompatible with production requirements. To support human in the analysis of performance indicators drifts, we propose in this paper a new generic and configurable diagnosis function based on production data collected and stored by the MES.

This paper focuses on proposing a set of potential causes which have most impact to performance indicators degradation and how to characterize them by analyzing MES data and recalling industrial expert's experience. In order to represent these cause and their failure modes on complex context of MES, probabilistic approaches such as Bayesian Network (BN) are well-suited techniques to build a diagnosis model based on MES data, the proposed approach enables online diagnosis for corrective actions after learning historical data phase.

The diagnosis objective and its targets are introduced in the next section. The problems are presented in section 3. Section 4 presents identified potential causes and propose of their characterized parameters according to our hypotheses of using unique MES database. Based that, BN diagnosis is proposed in section 5 to present identified variables relationship (failures modes and cause) to explain performance indicator drifts. At the end of this paper, the conclusion and future works are discussed.

2. MES DIAGNOSIS OBJECTIVES

Our global objective is to provide information on the origins of an OEE indicator's degradation (out of the threshold value) and to help making the best decision for all user categories (operator, team leader, supervisor and director) as described in Fig. 2. We focus on the difference between the measured OEE and the threshold value.

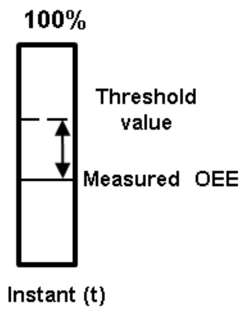


Figure 2. The distance between measured OEE and threshold value.

In order to achieve this purpose, we will begin with the definition of failure modes and sources causes in the following two subsections and basis to the determination of our chosen scientific problems for this paper.

2.1. Failure Modes

In fact, the OEE is consist of three component indicators: Availability, Performance and Quality which can be calculated on different time periods (hour, day, week, month, and year) and is described in Fig. 3 (Piétrac et al., 2011). In that way, the Availability, Performance and Quality drift are considered as three failure modes at the MES level.

According to Fig. 3, the Availability indicator correspond to the percentage of scheduled time the operation is available (with absence of downtime events in production process) to the total planned production time (with downtime losses). The Performance indicator represents the actual speed of the equipment as related to the design speed of the equipment. It is calculated as the number of total manufactured pieces divided to the product of operating time and theoretical cadence. Finally, the Quality indicator is the percentage of the

Availability *= Operating time/ Planned production time*

Performance *= Total pieces / Operating time x theoretical cadence*

Quality *= good pieces/ Total pieces*

OEE = Availability x Performance x Quality

Figure 3. Overall Equipment Effectiveness Formula.

good pieces to the total manufactured pieces.

In fact, the downtime losses which are the observed occurrences of events related to changes of state of the system are usually configured and measured via the OEE data inside MES data model. However they are considered as the secondary causes or intermediate causes but not the origin causes. Moreover, they do not support operators to explain Performance/ Quality indicator degradation on the analyze phase which become more difficult in incertitude and variability context. This present an example of the fact that historical MES data is not exploited to its full potential.

2.2. Sources Causes

In manufacturing context, these KPIs depend naturally on the performed result of production operation (operating time, quantity and quality of manufactured product). In our framework, the MES production station is generally represented by 5 important elements that impact the successful completion of production operations such as the operator (human factor), material, equipment, recipe and planning (as in Fig. 4). They are referred to in our paper as root causes. Any change of one of these elements may lead to changes of production process performance at both positively and negatively.

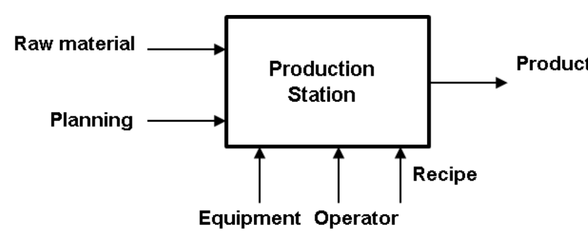


Figure 4. Important elements of MES production station.

Indeed, to achieve the global objective, MES diagnosis objective is to model and determine the relationship between failure modes and these causes as in Fig. 5. Naturally, it opens to the scientific problems of this paper and our contribution, presented in the next section.

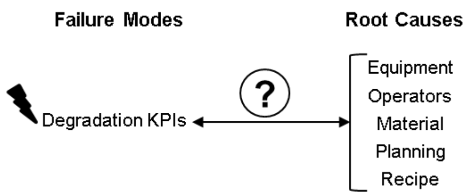


Figure 5. Diagnosis objective definition.

3. PROBLEMS

The scientific problems are organized in two different phases. The first phase is the identification of a set of variables to characterized root causes (such as the required qualification (skill) or experience of operators, work plan or health of equipment, quality of material, etc). They also represent the macroscopic results corresponding to the specific MES production station context. This is complemented with the determination of all the parameters that allow to characterize these variables on the different time horizons by using unique MES data.

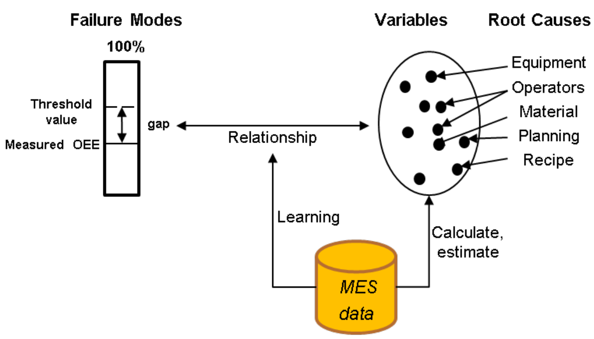


Figure 6. Problems.

In the second phase, we determine a well suited technique to exploit the rich MES database to model the relationship between these proposed failure modes and variables linked to root causes. Based on that, learning approach as BN is used to represent the cause-effect relationship and to evaluate the suspect level of variable by their probabilities after learning behavior of the production system.

4. FAILURE CAUSES OF MES

This section present a set of proposed potential causes variables around the sources causes such as the equipment (section 4.1), context production/planning (section 4.2), operator (section 4.3), recipe (section 4.4) and material (section 4.5) by calling for the industrial experience to help determine phase. This section also presents the characterization of these causes from the database of MES on the different time horizons (hour, day, week, month, and year). This section will be closed by the analysis diagnosis in order to model these relations of proposed causes and failures modes and localize the root causes by exploiting the historical production data MES.

4.1. Equipment Health index (EH)

The health situation of equipment is an important factor behind poor product quality (Abu-Samah et al., 2015; Nguyen et al., 2016). Therefore, we consider the Equipment Health (EH) as a important cause that impact the product quality and performance. In addition, the poor Equipment Health (EH) such as occurrence of breakdown event on equipment naturally impacts the availability indicator degradation (definition in OEE formula). So, how can we estimate the health of equipment level?

In order to respond this question, (Chen & Wu, 2007) evaluate the machine condition by observing the distribution of the machine parameters readings and comparing it against predefined machine specifications by Machine Capacity Index (MCI) indicator. The distance between observed and usual reading data parameters represents the deviation from its predefined target. It provides an easy reading of the equipment current operating condition. However, in this complex and ambiguous MES context, the number of equipment parameter is very large (such as in semi-conductor: millions sensors, and hundreds operations), it is difficult to compute the MCI, and worse in the case of incomplete data. Indeed, MCI is not well suited with the macroscopic level analysis (workstation) of our objective diagnosis.

In this complex context, (Bouaziz et al., 2011) proposed Equipment Health Factor (EHF) which is characterized by the probability of failure mode occurrence on equipment $EHF = \mathrm{P}(FM \backslash Observations)$. Based on the FMEA (Failure Modes and Effects Analysis) table, a BN model is structured with the failure modes, the root causes and parameters of equipment. By exploiting the historical production data, this model estimates the distribution probability to provide the last result EHF. This way, EHF can handle complex big data and even in the incomplete data case. Moreover, with the probability result (within 0 to 1), we can estimate several Equipment Health levels (ex: very low, low, normal, good and very good) to show better description of these causes. In addition, the quantity of necessary data to calculate EHF is available in the MES data model consist of historical production events, FMEA, maintenance events which was defined by the norm IEC 62264. Therefore, the estimation of EHF by MES data is possible.

However, (Bouaziz et al., 2011) proposed EHF as an indicator to predict the possibility of occurrence of FM on equipment at any instant (t). Therefore, in order to more pertinent and an active using in our case the Equipment Health (EH) should be calculated on different time periods (hour, day, week, month, and year) according to the purpose of users. We propose to estimate HE on any period (X) such as Fig. 7 by the average of all discretion instants of that period. In which X= $[t_0 \rightarrow t_n]$.

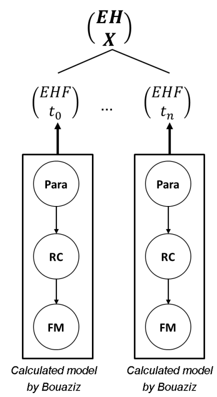


Figure 7. Estimated formula of EH by utilization EHF model.

According this figure, the HE on period (X) is estimated by:

$HE_{(X)}$= the average of $\{EHF_{(t0)}, ..., EHF_{(tn)}\}$

By this way, the HE is calculated on the different production periods (hours, day, week, month and year) depend on each of EHF(t). However, because of nature of learning phase, it needs a set of quantity sufficient data to calculate exactly probabilities (minimum 1000 learning samples). Therefore, it is necessary to know that the estimation EHF will be better on the long production periods such as week, month or year, several years .

4.2. Workload production Rate (WR) and mix production (Mix)

In highly competitive production environments which are constrained by rapid products and technologies changes, the equipment's can be often operated at their full capacities (Hubac & Zamai, 2013). Indeed, stressful work-flow where direct variability frequently and strongly influences the equipment performance, availability and quality product (Said et al., 2016). If that's happen in a long time, these will cause the performance degradation indicators which are considered failures modes in our work.

Causes		Formation	Note
WR	High	x> theoretical rate	$x = \frac{\text{manufactured product pieces}}{\text{net operating time}}$ (piece/h)
	Low	x≤ theoretical rate	
Mix	High	(number of change series > α) ∧ (change series time > β)	α: usual number of change series (discrete value) β: usual change series time (discrete value) α and β are predefine by experts
	Low	(number of change series ≤ α) ∨ (change series time ≤ β)	

Figure 8. The proposed estimated WR and Mix formation.

In order to estimate workload production rate (WR) and mix-production (Mix) level, (Duong et al., 2013) monitor the quantity of manufactured product, number type of product with quality product rate by a statistical technique where apply on a lot production in semi-conductor. By the same way, (Hohmann, 2011) propose to estimate the WR by comparing the quantity of manufactured product into a considered horizon production with maximal production capacity. Both of two are limited on their knowledge on the production system and especially their expertise.

Based on these works, we propose the estimated formation of these causes in the Fig. 8 with the similarity variables via OEE data inside MES data-model. These variables could be discretized by getting the average value in the preparing data phase and they are also measured on different time periods (1 hour, 8 hours, day, week, month, and year) to according to measured period of failure mode or OEE.

4.3. Human Level Experience (HLE), Fatigue Physical Human Level (FPHL) and Human Stress Level (HSL)

4.3.1. Human Level Experience (HLE)

Human behavior is an important and indispensable factor for the complex production process. However, they are potential source of disturbances because of lack of global vision on their decisions or the mistakes in their tasks. For example (Cacciabue, 2004; Nguyen et al., 2016) demonstrate ineffectiveness of human could cause production quality problem. In this context, we consider the Human Level Experience (HLE) as a failure cause impact to the performance and quality degradation.

In (Vemer et al., 2003), the HLE is defined as an acquisition of operational execution production knowledge over a long period of exercise activity on a workplace however this work does not give any specific way to measure HLE. To estimate this cause (HLE), we propose two way to estimate:

• Based on the definition of (Vemer et al., 2003), on the Fig. 9 we want to define the factors to estimate this cause. With α, x, β are the parameters which have determined by the learning phase the reference production data of one reference expert operator. This idea is very useful and well applied in the example as follows (in Fig. 10) with short calculation period (ex: 4 hours) consider a 6 years production data of several operator, each work shift is 4 hours with one operator and we get a table of correspond quality indicator consider as Performed level. With the determined , x, (assume), by reasoning via the equivalence ratio: $\frac{x}{\beta} = \frac{t}{\alpha} = \frac{beginning}{expert}$. With (t) is an instant in which estimate HLE.

In order to determine parameter α, x, β in this example, we analyse evolution of performance cadence of each operator on each 4 hours correspond to a work shift in Fig. 11. In

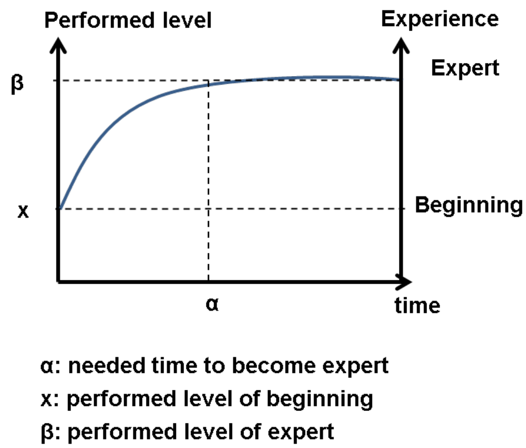


Figure 9. Estimate HLE by work time.

	Day 1	Day 2	...	Day 6	...
Morning (4h)	Operator A (Quality=80%)	Operator D (Quality=85%)	...	Operator A (Quality=86%)	...
Afternoon(4h)	Operator B (Quality=70%)	Operator B (Quality=84%)	...	Operator D (Quality=85%)	...
Night (4h)	Operator C (Quality=90%)	Operator F (Quality=76%)	...	Operator C (Quality=87%)	...

Figure 10. Example.

this case, the performance cadence could be calculated by the number of assemblage piece divided by performance period (4 hours). In fact, it is very obvious to find that the experience of operator is linear with the quantity of manufactured pieces. Then we determine the parameters:

+ x: correspond to cadence level of operator who is in first day on the workstation. Or we can estimate x as the lowest cadence level of operator on the first day of each operator.

+ β: correspond to cadence level of operator who has longest time on the workstation. Or we can estimate β as the highest cadence level of operator by observation their evolution.

+ α: correspond to duration whereas is intersect of most of evolution curves after they crossovers the β.

Based on this figure, the experience operator level of operator is defined according to his production time (t).

• Based on the reference model of personnel defined by IEC 62264 standard (page 92-93 of Part 1), we determine a set of parameter that impact to HLE factor: personnel, class personnel, personnel priority, class personnel priority, test specification priority, test specification result which were available on MES data. In addition, via the first idea and the definition of HLE, we must consider the indispensable temporal factor. Consequently, we consider on a model to present set of determined parameters (stochastic variables) and also temporal factor such as Dynamic Bayes Network (DBN). On context without hidden variable, we propose to estimate HLE by DBN in Fig. 12 base on its advantages of inference (diagnosis or prognosis) and also treatment incomplete data (Oliver & Horvitz, 2005):

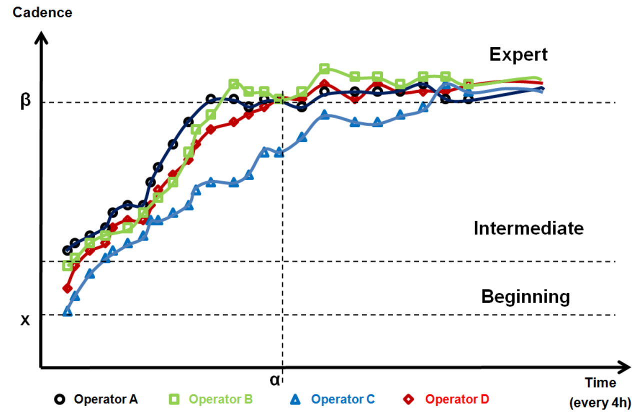


Figure 11. Example of experience curve.

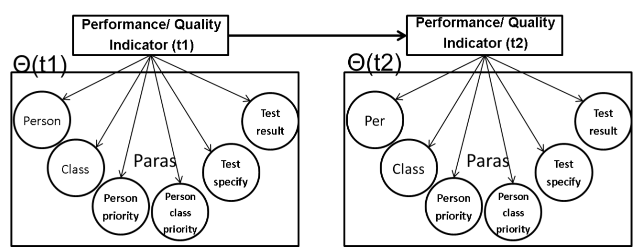


Figure 12. Estimate HLE by DBN.

This way, the HLE will be estimated by the probability of good performance indicators (OEE) knowing set of parameter on considered instant θ_{t2} and performance indicators on the past instant θ_{t1}.

$$HLE = \mathrm{P}(Perf_{\cdot\theta_{t2}} = good \setminus Paras_{\theta_{t2}}, Perf_{\cdot\theta_{t1}}).$$

In fact, this formula is well applied in (Duong et al., 2013) to estimate the Confidence of Reported Information CLFI. With probability within 0→1, we estimate HLE in plural moral values depending on users (beginner/intermediate and expert or good/normal and poor). With this DBN model, the HLE could be calculated in the different periods (hours, day, week, month or year) by different learning phase period. For example in Fig. 10 we estimate on a day period parameter Personnel ={ABC, DBF, ADC, FEB, etc} and Class personnel ={Class (ABC), Class (DBF), etc}. With two ideas to estimate for HLE, the choice is remaining to users according to their type of industrial purpose.

4.3.2. Fatigue Physical Human Level (FPHL) and Human Stress Level (HSL)

On the human factor, beside HLE cause, we also consider the Fatigue Physical Level and Stress Level as the failure causes which may increase the manufacturing risk. In fact, FPHL and HSL impact labor behavior which was symptom of the degradation performance indicators (Cacciabue, 2004). In this section, we present the FPHL and follows by the HSL. In literature, the FPHL is defined as the fatigue level appeared by physiological way of human body after a work period. In our manufacturing context, (Lan et al., 2003) propose to explain the factors impact to fatigue by a static model BN (Fig. 13). In reality, it is very difficult to estimate FPHL (or HSL) because each operator has different physical condition and different support capacity. Therefore, our framework does not contain physical condition and assume that work condition as a normal manufacturing environment.

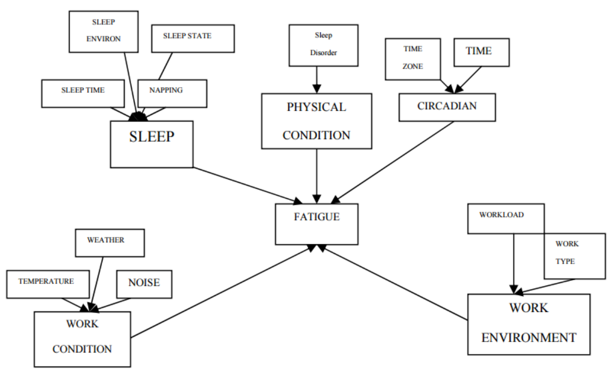


Figure 13. Static model BN to explain fatigue. (Lan et al., 2003)

a) In our manufacturing context, sleep time data and physical condition data will never exist in historical production data. Therefore, we propose to replace them to suit the MES production data using the break time for operator (ie: lunchtime, holiday, vacancy) when operators stop their works to relax. Based on that, we consider to estimate the work environment by WR and Mix variables. The relation between fatigue level and working time was proposed Pareto analysis by (McCulloch et al., 2007) as Fig. 14.
α and β were defined as the standard break and working time of operator. In France, α and β were proposed by (International Labour Conference, 2005) as 2 hours break/ 8 hours working on a working-day or 18 hours break/35 hours working on a week. We also measure them on the different production periods base on this portion. This way, we propose estimate the FPHL by the following formation:
- Fatigue= [($\beta \leq$ x) ∪ (y $\leq \alpha$)] ∩ [(Mix = high) ∩ (WR= high)].

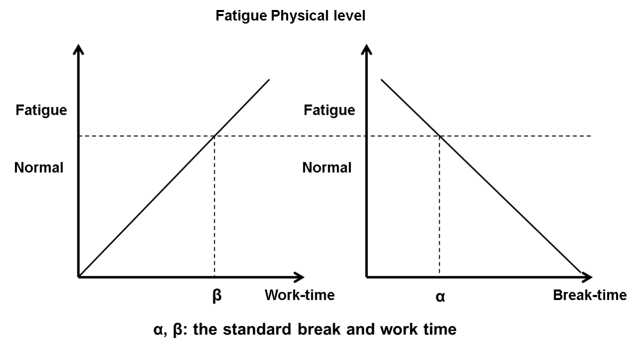


Figure 14. Relation between Fatigue Physical and break/working time.

- Normal= [(x $< \beta$) ∩ (y $> \alpha$)] ∪ [(Mix= low) ∩ (WR= low)].
In which, x and y as working and break time on a considered production period of a operator or even a team/group. In this case, they could be measured by taking the average of break/working time for each operator.

b) HSL on the other hand present the fatigue of human but in the side mental health: stress. The HSL is also defined in (Lan et al., 2003) as the effect of WR and Mix but also a complex level of work type.

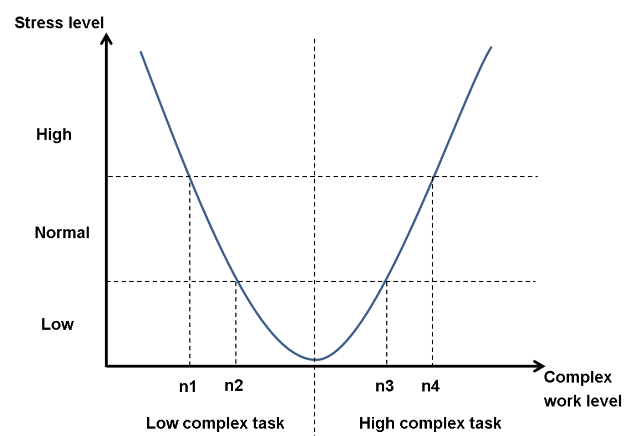


Figure 15. Relation between the stress of operator and complex work level.

Fig. 15 is proposed by (McCulloch et al., 2007) to present the relation between complex levels with HSL (stress of operator). In which $n1, n2, n3$ and $n4$ as the complex work levels. They could be estimated by Pareto analysis with the number of operation of equipment, because the more complex, the more number operation to execute. In another way, they could also be defined by expert by the operating mode of equipment. According to Fig. 15, the HSL can be estimated by the following formation:
- High= [z≥ $n4$ ∩ z≤ $n1$] ∪ [WR=high ∪ Mix=high]

- Normal= [$(n1$ <z< $n2) \cap (n3$ <z< $n4)$] ∪ [WR=low ∩ Mix=low].
- Low= [$n2 \leq$ z$\leq n3$] ∪ [WR=low ∪ Mix=low].
(With z is complex work level of considered equipment).

In reality, FPHL and HSL are two sides of operator capacity, they are involved together. When you feel stress because of work you will easily fatigue, and opposite.

4.4. Recipe Change (RC)

The concept of recipe is widely used in the manufacturing systems. The recipe is often developed by the R&D and qualified on equipment before using on production system. The control recipe often consists: the production schedule, the multi-purpose plant description and executable control recipe is presented in (Genrich et al., 1994). In reality, the oversights during the design phase or transition phase lead to poor product quality (Nguyen et al., 2016). At that transition phase, the change recipe process which represent the change process of number parameters on machine can lead to poor product quality (nonconformity piece). Moreover, by Pareto analysis (Duong et al., 2013) demonstrate that the RC have strong impact to the product quality through increasing the equipment drift. Therefore, we consider in this case the RC is an important cause that produce the quality degradation.

Via (Genrich et al., 1994), we define a set of parameters which present the most impact of change recipe: cadence (rate), installation machine time, number procedure, procedure version, and product type (Meyer et al., 2009). Similar to inference in section 3.3, we model process change recipe by DBN because of presentation temporal factor as Fig. 16.

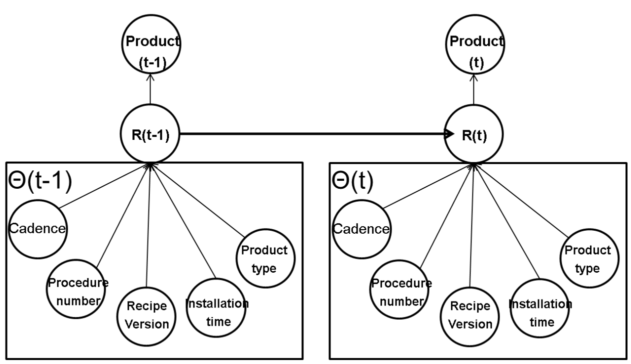


Figure 16. Estimate RC by RBD.

This way, the RC will be estimated by the probability of good piece (price conformity) knowing set of parameter on considered instant θ_t and quality of product on the past instant θ_{t-1}. With a probability within 0→1, we estimate RC in plural moral values depend on users. In this work, we consider RC with 3 moral values low, normal and high. Such as EHF, the RC can be estimated on the different production periods production data.

4.5. Raw Material Quality (MQ)

Material is one of the important external factor impact to production operation, especially hazards of the characteristics of the raw material or quality of the raw material. In fact, raw material quality has a big effect on the physical product characteristic that lead to poor quality product from a poor material quality (Ashori & Nourbakhsh, 2008). Hence, correspond to our macroscopic level analysis of our work, our objective specify the impact level between material quality and a degradation of OEE indicators on a considered production period.

Inside MES data model, the material model is defined by the material lot/class, material definition, properties and material specification (IEC 62264). Under which, material quality is perhaps characterized by the QA (Quality Assurance) test result that represents report of good characteristic of a material lot (ex: pass or fail). That way, a statistical technique is used for estimating this cause from historical production data in following formation. With probability, we considered this cause in the plural morale (good/normal/bad ...).

$MQ = \mathrm{P}(Product = bad \setminus QA - test - result = fail).$

This way, QA can be estimated on the different production periods.

5. DIAGNOSIS APPROACH BASED ON BAYESIAN NETWORK

Through section 2 and section 4 we defined the failure modes and the potential causes in our case of MES. But how can we model this cause and failure mode to show their relation in order to specify the failure cause?

Based on a rich MES database, a method which can exploit the historical production data to achieve our objective diagnosis will be the focus. Moreover, in the dynamic context, that kind of method is well-suited with the variation product, operators, and recipe, etc. in production context. In addition, through the determined failures causes space we should focus in how to evaluate suspected causes level between them and how to describe exactly and dynamic their relations.

All these needs lead us to the probabilistic approaches. These approaches can be performed without understanding the underlying structure of a production system (Bouaziz et al., 2011). Among the probabilistic approaches, we incline toward the Bayesian Network (BN) which is widely used to identify a graphical structure model that describes relationships between variables in production system (Weber et al., 2012). Moreover, its conditional probabilities will be calculated by a learning phase to provide the risk priorities of causes and support corrective maintenance decisions (Nguyen et al., 2016). However, a difficulty of BN is to identify the graphical structure in complexity context of manufacturing system. Identification of this graphical structure is often be performed by experts or can be learned from the data using many algorithms such as K2, PC or Tabu. Consequently, our contribution in section 5 define the relation between these proposed causes and failures modes by transformation the relations of their characteristics parameters in meta data model MES and propose suite cased BN model.

5.1. Established BN model

In order to establish BN model, 3 principal phases: identification variables, determine relation between variables and determine BN probability law as proposed in (Bouaziz et al., 2011). Following our study in these previous sections, our BN model is developed basic on the set of proposed variables that represented relation causes-effects corresponds to nodes in the static Bayesian model with:

- Failure modes: As we mention, three failure modes are correspond to three component OEE indicators: Availability, Performance and Quality. They could be measure likes continuous and discrete quantitative variables on different production periods. In that case, we assume that the state of these variables takes three values (good, normal and bad) by their average value to indicate level of indicators variances.
- Failure causes: these variables represent potential causes in MES as mention in section 4. Their values have calculated by collecting their parameters from historical production data. Therefore, they are often taken plural moralities described under the formalization phase or discrete values (good, normal or bad) by their probabilities.

Second phase, the task of determination relations between the identified variables is often performed into two steps. First, we define relations between nodes of the different families (causes and failure modes) and second, we define links between the different variables nodes of each group.

In our cases, the first step is performed at definition phase of each cause variable. For example, Equipment Health (EH) is characterized by the failure modes and parameters of machine, so it impact to product quality (Bouaziz et al., 2011) and machine availability or availability and quality indicators. In addition, Workload Rate (WR) and product mix (Mix) are represented by the number manufactured pieces, number series changes or series changes times of production that impact to variant of performance and quality indicators (norm EF 60). Besides that, the operator mistakes lead to the failures on products or production performance (Cacciabue, 2004; Nguyen et al., 2016). Other means implies these cause consequence relationships HLE, FPHL, HSL $\rightarrow$ Performance and Quality indicators. Therewith, basic on (Abu-Samah et al., 2015) and (Ashori & Nourbakhsh, 2008), Change recipe (RC) and Material Quality (MQ) may lead to poor product quality.

Second, in order to draw links between the different variables nodes of each group, there is interface between human factor, control recipe and equipment in context of a successful production operation, presented in (Nguyen et al., 2016). Under which, the oversights of human factors may lead to failure on equipment and furthermore, the failure in design step or change phase of recipe are also too. Moreover, they in fact are provided by transformation relations between their parameters in standard data model MES. Indeed, failures operator sources (HLE, FPHL, HSL) and failure recipe source (RC) changes their characterized parameters likes work-plan times, operations resources or production parameters who can cause the variant of production data (events, measurement or quality data) of equipment. In addition, there is also the causes-consequence relationship in which Mix and WR affect to FPHL and HSL as their estimated formulas in section 4.

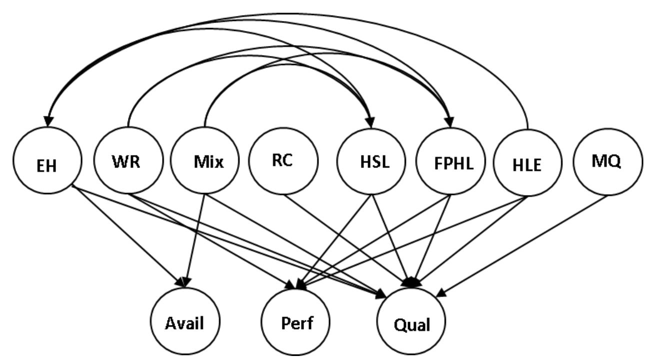


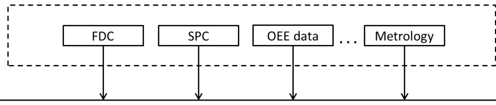
Figure 17. Causal graph BN of identified nodes.

Based on the analysis of relationships between failure source and identified failure modes, we propose in this section the causal graph as presented in Fig. 17.

We take careful note these BN properties in our case as no loops, no hidden node or absence of temporal properties. Therefore, we consider Tree Augmented Naive Bayes (TAN) to represent failure mode class has no child while each cause node has maybe the parent class or a child. We take also a strong assumption in this paper that all nodes take discrete value and are observed to simply the practice phase.

5.2. Required data collection

The learning BN phase needs to have a table where each column corresponds to a node with each line records on a time periods (either hour, day, week, month or year). In this phase, we call to the industrial experience to determine the appropriated period to estimate these variables according to type of industrial product, for example a week for semiconductor or months for a medical product. Following we present in Fig. 18 a collected data table example of AIP-RAO (education manufacturing system) with day line production record.



Collected production data analyzing and processing													
ID	Date_START	Date_End	Availability	Perfor mance	Quality	EH	WR	Mix	HLE	FPHL	MHHL	RC	MQ
78	9/12/2010 07:30:15	9/12/2010 22:15:23	Good	Normal	Good	Bad	Low	High	Good	Fatigue	Normal	Normal	Good
79	10/12/2010 08:15:21	10/12/2010 22:00:12	Bad	Normal	Good	Normal	High	Low	Normal	Fatigue	High	Low	Bad
80	11/12/2010 07:45:14	12/12/2010 00:09:48	Good	Good	Good	Bad	Low	Low	good	Normal	Low	Normal	Good
...	...	...	...	...	...	...	...	...	...	...	...	...	...

Figure 18. Production data table.

The production data table above is a collected production data table example in 6 month at station 4 from MES-AIP data. In which, the Availability, Performance and Quality indicators are calculated on day, we assumed to estimate them by comparing with an objective indicators which is usually defined by an industrial expert. These causes variables have estimated according to their formula mention in section 4 on same day periods.

5.3. Diagnosis approach Execution

In this section, we propose an algorithm to diagnosis from the structure presented in Fig. 17, based on computing conditional probabilities of this model. We started with data analysis to identify the relationships and impact on these performances indicators of MES through the experience and knowledge of experts who built the Bayesian Network structure from the instantiated MES production data model as described in Fig. 19. From now, the result of this diagnosis model will be provided online after learning phase a considered production data at where needs to explain the degradation performance indicators.

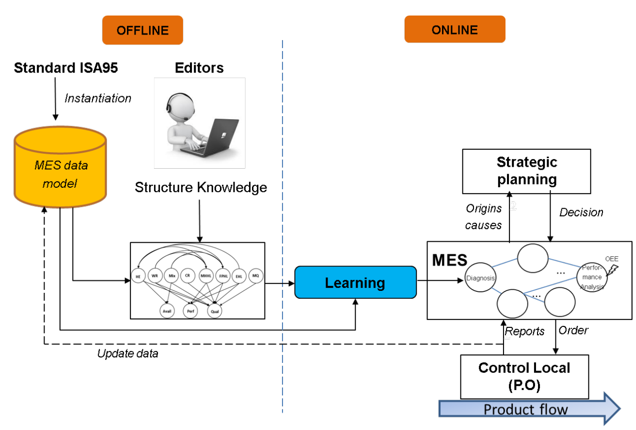


Figure 19. The diagnosis approach execution.

By estimating the conditional probabilities of each component (causes variables) with failure modes variables presented in Fig 20. The final decision will be made by the suspected

Calculated Algorithm

Input: *Instantiation bdd MES*
Treatment: From given MES data

- *Determine measured OEE components and estimated causes variables periods.*
- *Prepare the processing data to estimate each of variables (mention in section 3).*

Learning:

- *Calculate « priori probabilities » P(Avail), P(Perf) and P(Qual) corresponds failure modes: Availability, Performance and quality.*
- *Calculate conditional probabilities of each cause together: P(EH|RC,HLE,FPHL,MHHL); P(FPHL|Mix,WR), P(MHHL|Mix,WR)…*
- *Calculate conditional probabilities each cause and failure modes: P(Causes/Failure modes): P(EH=bad |Avail=bad), P(EH=bad| Qual=bad), P(EH | Avail, Qual) …*

Output:

➢ *Probabilities table P(Causes/Failures modes) to make decision*

Figure 20. Proposed approach algorithm.

components priorities.

Based on process data and the experts knowledge, we are now able to provide necessary information to explain the degradation function of MES though by 3 component performance indicators: Availability, Performance and Quality. This result is a set of value between 0 and 1 that extends the first phase of on-line diagnosis inference in MES level.

In addition, the diagnosis process is performed based on a BN model in which a rich MES database has not deployed just to analysis performance phase but also to characteristic the MES potential causes even if in incomplete data case.

6. CONCLUSION

In this paper, we proposed at first the macroscopic analysis corresponding to the specific MES level to provide maximum information of the origins of an OEE degradation. Through an analysis of MES standard data model who defines all the production data on the basis of the ISA 95, we determined a set of potential causes that may impact the successful completion of production operations such as the operator stress, quality material, equipment or change recipe, etc. This phase has not only presented to identify the observable causes via standard data model of MES but also all the parameters that allow to characterize these causes from the database of MES on the different time horizons. Based on that, we used a BN model to generate the associated probabilities on the identified structure BN to evaluate the suspect level of each proposed potential fault origins. From that, the diagnosis results support quickly and right decision-making for corrective maintenance activities in MES level. In next time, we focus on testing the model on data collected from an experimental manufacturing system in AIP-RAO. And next, we study to validate this fault diagnosis approach in a real data of an industrial case MES to improve optimization and accurate this model.

ACKNOWLEDGMENT

The authors gratefully acknowledge AIP-RAO of Lyon for their support and provision of MES experimental system for case study. The authors also acknowledge region Rhone Alpes for ongoing Research.

REFERENCES

Abu-Samah, A., Shahzad, M. K., Zamaï, E., & Hubac, S. (2015). Effective maintenance by reducing failure-cause misdiagnosis in semiconductor industry (si). *PHM Journal*.

Ashori, A., & Nourbakhsh, A. (2008). Effect of press cycle time and resin content on physical and mechanical properties of particleboard panels made from the underutilized low-quality raw materials. *Industrial crops and products*, *28*(2), 225–230.

Bouaziz, M. F., Zamai, E., Duvivier, F., & Hubac, S. (2011). Dependability of complex semiconductor systems: Learning bayesian networks for decision support. In *Dependable control of discrete systems (dcds), 2011 3rd international workshop on* (pp. 7–12).

Cacciabue, P. C. (2004). Human error risk management for engineering systems: a methodology for design, safety assessment, accident investigation and training. *Reliability Engineering & System Safety*, *83*(2), 229–240.

Chen, A., & Wu, G. (2007). Real-time health prognosis and dynamic preventive maintenance policy for equipment under aging markovian deterioration. *International Journal of Production Research*, *45*(15), 3351–3379.

Duong, Q.-B., Zamai, E., & Tran-Dinh, K.-Q. (2013). Confidence estimation of feedback information for logicdiagnosis. *Engineering Applications of Artificial Intelligence*, *26*(3), 1149–1161.

Genrich, H. J., Hanisch, H.-M., & Wöllhaf, K. (1994). Verification of recipe-based control procedures by means of predicate/transition nets. In *Application and theory of petri nets 1994* (pp. 278–297). Springer.

Hohmann, C. (2011). *Techniques de productivité: Comment gagner des points de performance-pour les managers et les encadrants*. Editions Eyrolles.

Hubac, S., & Zamai, E. (2013). Politiques de maintenance equipment en flux de production stressantequipment maintenance policy in stressed manufacturing flow (technology or product). *Edition TI (Technique de lingenieur)[AG 3535]*.

IEC, N. (2003). 62264-1. *Intégration du système de conduite d'entreprise-Partie*, *1*.

International Labour Conference, r. s. (2005). *Hours of work : from fixed to flexible?* International Labour Office.

Lan, P., Ji, Q., & Looney, C. G. (2003). Non-intrusive real time human fatigue modelling and monitoring. In *Proceedings of the human factors and ergonomics society annual meeting* (Vol. 47, pp. 311–315).

McCulloch, K., Baker, A., Ferguson, S., Fletcher, A., & Dawson, D. (2007). Developing and implementing a fatigue risk management system. *Transport Canada: Canada*.

Meyer, H., Fuchs, F., & Thiel, K. (2009). *Manufacturing execution systems (mes): Optimal design, planning, and deployment: Optimal design, planning, and deployment*. McGraw Hill Professional.

Nguyen, D. T., Duong, Q. B., Zamai, E., & Shahzad, M. K. (2016). Fault diagnosis for the complex manufacturing system. *Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability*, 1748006X15623089.

Oliver, N., & Horvitz, E. (2005). A comparison of hmms and dynamic bayesian networks for recognizing office activities. In *User modeling 2005* (pp. 199–209). Springer.

Piétrac, L., Lelevé, A., & Henry, S. (2011). On the use of sysml for manufacturing execution system design. In *Etfa 2011* (pp. 1–8).

Said, A. B., Shahzad, M. K., Zamai, E., Hubac, S., & Tollenaere, M. (2016). Experts knowledge renewal and maintenance actions effectiveness in high-mix low-volume industries, using bayesian approach. *Cognition, Technology & Work*, *18*(1), 193–213.

Vemer, L., Oleynikov, D., Holtmann, S., Haider, H., & Zhukov, L. (2003). Measurements of the level of surgical expertise using flight path analysis from da vinci robotic surgical system. *Medicine Meets Virtual Reality 11: NextMed: Health Horizon*, *94*, 373.

Weber, P., Medina-Oliva, G., Simon, C., & Iung, B. (2012). Overview on bayesian networks applications for dependability, risk analysis and maintenance areas. *Engineering Applications of Artificial Intelligence*, *25*(4), 671–682.